

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 29, 2003, 20:49:12 ; Search time 490.002 Seconds

(without alignments)
8229.920 Million cell updates/sec

Title: US-09-988-971-1_COPY_694_942

Perfect score: 249

Sequence: 1 tggctgctatgagggccctgag.....agggccctgctgagccatrac 249

Scoring table: IDENTITY NUC

Gapop 10.0, Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
1: EST.*
2: em_estba.*
3: em_estum.*
4: em_estin.*
5: em_estmu.*
6: em_estov.*
7: em_estro.*
8: em_hic.*
9: gb_est1.*
10: gb_est2.*
11: gb_hic.*
12: gb_est3.*
13: gb_est4.*
14: gb_est5.*
15: em_estfun.*
16: em_estom.*
17: gb_gss.*
18: em_gss_hum.*
19: em_gss_inv.*
20: em_gss_pln.*
21: em_gss_vrt.*
22: em_gss_fun.*
23: em_gss_mam.*
24: em_gss_mus.*
25: em_gss_other.*
26: em_gss_pro.*
27: em_gss_rtd.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	247.4	99.4	1002	14	B0052308
2	238	95.6	1069	14	B0052468
3	236	94.8	1020	14	B0054281
4	213.8	85.9	926	11	AK020837
5	209.6	84.2	877	9	AL541041
6	163.4	65.6	377	9	AA959151

7	147	59.0	566	12	BG284179
8	119.6	48.0	660	10	BB635615
9	107.2	43.1	986	14	B0054265
10	100.2	40.2	976	9	AL543079
11	100.2	40.2	1118	14	B0056298
12	99.2	39.8	627	10	BB619854
13	98.6	39.6	528	13	B1681002
14	98.6	39.6	521	13	B1345674
15	97	39.6	477	10	BE268290
16	97	39.0	783	13	B1820532
17	97	39.0	830	12	BG756739
18	97	39.0	996	14	BM919471
19	97	39.0	1055	14	BM918144
20	97	39.0	1192	13	BM545575
21	96.6	38.8	803	9	AU125413
22	95	38.2	800	13	B1839016
23	95	38.2	824	12	BF129084
24	95	38.2	998	12	BF129162
25	93	37.3	783	10	BE396636
26	90.2	36.2	390	10	AM487302
27	90	36.1	403	13	B1541790
28	89.8	36.1	916	12	BE745984
29	88.2	35.4	795	9	AL557199
30	88.2	35.4	1028	14	BM817288
31	88	35.3	558	13	B1813104
32	88	35.3	578	13	B1681042
33	87.6	35.2	891	14	B0051595
34	87.6	35.2	954	14	BQ711423
35	87.6	35.2	995	9	AL548206
36	87.6	35.2	1066	14	B0052951
37	86	34.5	862	13	B1819198
38	86	34.5	921	12	BG743960
39	85	34.1	1026	14	BQ662291
40	84.4	33.9	825	9	AL559108
41	84.2	33.8	2814	11	BC003253
42	82.6	33.2	286	14	R74940
43	82.6	33.2	539	13	B1900513
44	82.6	33.2	618	10	AW228890
45	82.6	33.2	693	14	BM949966

ALIGNMENTS

RESULT 1
B0052308
LOCUS
DEFINITION AGENCOURT 6868571 NIH_MGC_106 Homo sapiens CDNA clone IMAGE:5933542
5', mRNA sequence.
ACCESSION B0052308
VERSION B0052308.1 GI:19811648
KEYWORDS
SOURCE
ORGANISM
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Carnivora; Hominoidea; Homo.

REFERENCE
1 (bases 1 to 1002)
NITH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Dr. Daniel McGivern, DBS/NCI
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
Plate: L1CM2118 row: d column: 23
High quality sequence stop: 670.
Location/Qualifiers
1..1002

FEATURES
source

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5933542"
/clone_1b="NIH_MGC_106"
/issue_type="natural killer cells, cell line"
/lab_host="DH10B (phage-resistant)"
/issue="Organ: blood; Vector: pOTB7; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCAAGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."

```

BASE COUNT 221 a 296 c 288 g 197 t
ORIGIN

Query Match 99.4%; Score 247.4; DB 14; Length 1002;
Best Local Similarity 99.6%; Pred. No. 8.5e-58;
Matches 248; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY 1 TGGCTGTATGAGGGCTTCAAGAGGAGAAAGAGAGAACTGTTTACTCTGGGAC 60
DB 375 TGGCTGTATGAGGGCTTCAAGAGGAGAAAGAGAGAACTGTTTACTCTGGGAC 434
QY 61 CTTGAGAGGGGCTTCTCATCCGGAGAGCCAGACAGAGAGGCTTTACTCTGTCA 120
DB 435 CTTGAGAGGGGCTTCTCATCCGGAGAGCCAGACAGAGAGGCTTTACTCTGTCA 494
QY 121 GTCCGCTTCAAGCCGCTTCAATCTCGGAGCCGATCAAGATCAAGATCAAGCTT 180
DB 495 GTCCGCTTCAAGCCGCTTCAATCTCGGAGCCGATCAAGATCAAGATCAAGCTT 554
QY 181 GACAAATGGCTGTATGATCTCAACCGGCTTCACTTCCCTTCACTCCAGGCTGTG 240
DB 555 GACAAATGGCTGTATGATCTCAACCGGCTTCACTTCCCTTCACTCCAGGCTGTG 614
QY 241 GACCAATAC 249
DB 615 GACCAATAC 623

```

RESULT 2 1069 bp mRNA linear EST 29-MAR-2002
LOCUS B0052468
DEFINITION AGENCOURT 6868422 NIH_MGC_106 Homo sapiens cDNA clone IMAGE:5933772
5', mRNA sequence.
ACCESSION B0052468
VERSION B0052468.1 GI:19811808
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE NIH-MGC http://mgc.nci.nih.gov/
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: Dr. Daniel McVicar, DBS/NCI
cDNA Library Preparation: Rubin Laboratory
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: L1062118 row: n column: 13
High quality sequence stop: 681.
Location/Qualifiers
1. 1069
/organism="Homo sapiens"
/db_xref="taxon:9606"

FEATURES
source

```

/clone="IMAGE:5933772"
/clone_1b="NIH_MGC_106"
/issue_type="natural killer cells, cell line"
/lab_host="DH10B (phage-resistant)"
/issue="Organ: blood; Vector: pOTB7; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCAAGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."

```

BASE COUNT 230 a 328 c 300 g 205 t 6 others
ORIGIN

Query Match 95.6%; Score 238; DB 14; Length 1069;
Best Local Similarity 100.0%; Pred. No. 3.4e-55;
Matches 238; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 TGGCTGTATGAGGGCTTCAAGAGGAGAAAGAGAGAACTGTTTACTCTGGGAC 60
DB 368 TGGCTGTATGAGGGCTTCAAGAGGAGAAAGAGAGAACTGTTTACTCTGGGAC 427
QY 61 CTTGAGAGGGGCTTCTCATCCGGAGAGCCAGACAGAGAGGCTTTACTCTGTCA 120
DB 428 CTTGAGAGGGGCTTCTCATCCGGAGAGCCAGACAGAGAGGCTTTACTCTGTCA 487
QY 121 GTCCGCTTCAAGCCGCTTCAATCTCGGAGCCGATCAAGATCAAGATCAAGCTT 180
DB 488 GTCCGCTTCAAGCCGCTTCAATCTCGGAGCCGATCAAGATCAAGATCAAGCTT 547
QY 181 GACAAATGGCTGTATGATCTCAACCGGCTTCACTTCCCTTCACTCCAGGCTGTG 238
DB 548 GACAAATGGCTGTATGATCTCAACCGGCTTCACTTCCCTTCACTCCAGGCTGTG 605

```

RESULT 3 1020 bp mRNA linear EST 29-MAR-2002
LOCUS B0054281
DEFINITION AGENCOURT 6830234 NIH_MGC_106 Homo sapiens cDNA clone IMAGE:5936362
5', mRNA sequence.
ACCESSION B0054281
VERSION B0054281.1 GI:19813621
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE NIH-MGC http://mgc.nci.nih.gov/
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: Dr. Daniel McVicar, DBS/NCI
cDNA Library Preparation: Rubin Laboratory
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: L1062125 row: j column: 11
High quality sequence stop: 556.
Location/Qualifiers
1. 1020
/organism="Homo sapiens"
/db_xref="taxon:9606"

FEATURES
source

1. 1020
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5936362"
/clone_1b="NIH_MGC_106"
/issue_type="natural killer cells, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: blood; Vector: pOTB7; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned

FEATURES
source

Location/Qualifiers
1. .926
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="PANTOM DB:A930009E21"
/db_xref="MCD:MGI:1911678"
/db_xref="taxon:10090"
/clone="A930009E21"
/issue_type="retina"
/clone_id="R1XEN full-length enriched mouse cDNA library"
/dev_stage="adult"
128. .667
/note="data source:SPRR, source key:Q60898, evidence:ISS
putative
similar to SRC-LIKE ADAPTER PROTEIN"
/codon_start=1
/protein_id="BAB32223.1"
/db_xref="GI:12861543"
/db_xref="MCD:MGI:1925049"
/translation="MPSYVAVKVAHGMVLEGLSREKAEELLLPDPGAFILRESQT
RRGYSLSVLRSPASMDRIHRYRIORLNDGMVYISPLRTPSLFLAVERYSLDGI
CCPLREPCVLQKLGPLRGKDPPTPTVPTSLMVKKLDRLSLFLAVERYSLSLSEG
LRESLSYSLSLADPPLDDA"

polyA_signal
907. .912
/note="putative"

polyA_site
926
/note="putative"

BASE COUNT 209 a 269 c 263 g 185 t

ORIGIN

Query Match
Best Local Similarity 91.2%; Score 213.8; DB 11; Length 926;
Matches 227; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

QY 1 TGGCTGTATGAGGGCCTTGACAGAGGAGAAAGCAGAGAACTGCTGTTTACTCTGGGAAAC 60
Db 164 TGGCTGTATGAGGGCCTTGACAGAGGAGAAAGCAGAGAACTGCTGTTTACTCTGGGAAAC 223

QY 61 CCGTGAAGGGGCTTCTCTATCCGGGAGAGCCAGACAGAGAGGCTCTTACTCTGTCACA 120
Db 224 CCGGAGAGGGCTTCTCTATCCGGGAGAGCCAGACAGAGAGGCTCTTACTCTGTCACC 283

QY 121 GTCCGCTCAGCCGCCCTTGATCTCTGGGACCGGATCAACAATAAGATCACTGCTT 180
Db 284 GTCCGACTCAGCCGCCCTTGATCTCTGGGACCGGATCAACAATAAGATCACTGCTT 343

QY 181 GACCATGGCTGGCTGTACATCAACCGGCTCACTCCCTCACTCCAGCCCGCTGGTG 240
Db 344 GACCATGGCTGGCTGTACATCACTCCGCTCACTCCCTCACTCCAGCCCGCTGGTG 403

QY 241 GACCATTAC 249
Db 404 GACCATTAC 412

RESULT 5
AL541041 877 bp mRNA linear EST 16-FEB-2001
LOCUS AL541041 LIT_FL002_P11 Homo sapiens cDNA clone CS0DE005YK23 5 prime
DEFINITION ' mRNA sequence.
ACCESSION AL541041
VERSION AL541041.1 GI:12871733
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Genoscope - Centre National de Sequencage

BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr.

FEATURES
source

Location/Qualifiers
1. .877
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0DE005YK23"
/clone_id="LIT_FL002_P11"
/lab_host="DH10B"
/note="Organ: placenta; Vector: pCMVSPORT 6; 1st strand
cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-stranded cDNA was digested with Not I
and cloned into the Not I and Eco RV sites of the
pCMVSPORT 6 vector. Library was constructed by Life
Technologies. Contact: Feng Liang Life Technologies, a
division of Invitrogen 9800 Medical Center Drive Rockville
, Maryland 20850, USA Fax: (1) 301 610 8371 Email:
http://fulllength.invitrogen.com"

BASE COUNT 192 a 262 c 251 g 170 t 2 others

ORIGIN

Query Match
Best Local Similarity 98.6%; Score 209.6; DB 9; Length 877;
Matches 209; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 TGGCTGTATGAGGGCCTTGACAGAGGAGAAAGCAGAGAACTGCTGTTTACTCTGGGAAAC 60
Db 666 TGGCTGTATGAGGGCCTTGACAGAGGAGAAAGCAGAGAACTGCTGTTTACTCTGGGAAAC 725

QY 61 CCGTGAAGGGGCTTCTCTATCCGGGAGAGCCAGACAGAGAGGCTCTTACTCTGTCTCA 120
Db 726 CCGTGAAGGGGCTTCTCTATCCGGGAGAGCCAGACAGAGAGGCTCTTACTCTGTCTCA 785

QY 121 GTCCGCTCAGCCGCCCTTGATCTCTGGGACCGGATCAACAATAAGATCACTGCTT 180
Db 786 GTCCGCTCAGCCGCCCTTGATCTCTGGGACCGGATCAACAATAAGATCACTGCTT 845

QY 181 GACCATGGCTGGCTGTACATCTCAGCCGCCCT 212
Db 846 GACCATGGCTGGCTGTACATCTCAGCCGCCCT 877

RESULT 6
AA959151 377 bp mRNA linear EST 08-MAY-1998
LOCUS AA959151
DEFINITION v51906.r1 Soares thymus 2BMT Mus musculus cDNA clone
IMAGE:1330042 5' similar to TR:013239 Q13239 PUTATIVE SRC-LIKE
ADAPTER PROTEIN ; mRNA sequence.
ACCESSION AA959151
VERSION AA959151.1 GI:3124344
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
AUTHORS Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,
Geisler, S., Kucada, T., Lacy, M., Le, M., Martin, J., Morris, B.,
Schellenberg, K., Steppe, M., Tan, F., Underwood, K., Moore, B.,
Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and
Waterston, R.
TITLE The WashU-HMI Mouse EST Project
JOURNAL Unpublished (1996)
COMMENT Contact: Marra M/Mouse EST Project
WashU-HMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LIND; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.

BP 191 91006 ENRY cedex - France
Email: seqref@genoscope.cns.fr, web : www.genoscope.cns.fr.
Location/Qualifiers

FEATURES

SOURCE

1. 976
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS001002YL19"
/clone_1lb="UTL NPL006 PL2"
/tissue_type="Placenta"
/note="Vector: PCWVSORT 6; Site 1: NotI; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the PCWVSORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Peng Liang Life Technologies, Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
Email : liliang@litech.com URL :
http://fulllength.invitrogen.com"

BASE COUNT 237 a 280 c 300 g 158 t 1 others
ORIGIN

Query Match 40.2%; Score 100.2; DB 9; Length 976;
Best Local Similarity 62.7%; Pred. No. 3.7e-17;

Matches 156; Conservative 0; Mismatches 93; Indels 0; Gaps 0;

QY 1 TGGCTGTATGAGGCGCTTGAGCAGGAGAAAGCAGAGAACTGCTGTTTACTCTGGGAAC 60
DB 586 TGGTTTTCAGAGGCGATCAGCCGAGAGAGCAGAGGCCCACTGCTGCTCCGGCAAC 645
QY 61 CTTGAGAGGGGCTTCTCATCCGAGAGCCAGACCAAGAGAGGCTTACTCTCTGTCA 120
DB 646 ATGCTGGCTCTTATGATTCGGGATACCGAGACCACTAAAGAGCTACTCTTTGTCC 705
QY 121 GTCCGCTTACGCCCTTGCATCTCGGAGCCGATCAGACATCAAGATCCACTGCTT 180
DB 706 GTCCGAGACTAGACCTTCGAGAGAGATACCGTAAACATTACAGATCCGAGCCCTG 765
QY 181 GACAAATGCTGCTGTATCATCTCAGCCGCTCAGCCTTCCCTCATCTCAGAGCCCTGGT 240
DB 766 GACAAAGGGGGCTTCTACATATCCCCCAAGACACTTTCAGCACTGTGAGAGCTGGT 825
QY 241 GACCAATTAC 249
DB 826 GACCACTAC 834

RESULT 11

BQ056298 1118 bp mRNA linear EST 29-MAR-2002
LOCUS BQ056298
DEFINITION AGENCOURT_6773554 NIH_MGC_99 Homo sapiens cDNA clone IMAGE:5808612
5', mRNA sequence.
ACCESSION BQ056298
VERSION BQ056298.1 GI:19915638
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE

1 (bases 1 to 1118)
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
NIH-MGC http://mgi.nci.nih.gov/.

AUTHORS

Unpublished (1999)
National Institutes of Health, Mammalian Gene Collection (MGC)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: Lou Staudt

TITLE

CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.lnl.gov
Plate: L1CM2052 row: 0 column: 13

High quality sequence stop: 489.
Location/Qualifiers

FEATURES

SOURCE

1. 1118
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5808612"
/clone_1lb="NIH_MGC_99"
/tissue_type="Lymphoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: lymph. Vector: pOT87; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGACGAG(G). Size-selected >500bp for average insert size 1 kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC library."

BASE COUNT 277 a 323 c 324 g 194 t
ORIGIN

Query Match 40.2%; Score 100.2; DB 14; Length 1118;
Best Local Similarity 62.7%; Pred. No. 3.9e-17;

Matches 156; Conservative 0; Mismatches 93; Indels 0; Gaps 0;

QY 1 TGGCTGTATGAGGCGCTTGAGCAGGAGAAAGCAGAGAACTCTGTTTACTCTGGGAAC 60
DB 119 TGGTTTTCAGAGGCGATCAGCCGAGAGAGCAGAGGCCCACTGCTGCTCCGGCAAC 178
QY 61 CTTGAGAGGGGCTTCTCATCCGAGAGCCAGACCAAGAGGCTTACTCTGTCTCA 120
DB 179 ATGCTGGCTCTTATGATTCGGGATACCGAGACCACTAAAGAGCTACTCTTTGTCC 238
QY 121 GTCCGCTTACGCCCTTGCATCTCGGAGCCGATCAGACCAAGATCCACTGCTT 180
DB 239 GTCCGAGACTAGACCTTCGAGAGAGATACCGTAAACATTACAGATCCGAGCCCTG 298
QY 181 GACAAATGCTGCTGTATCATCTCAGCCGCTCAGCCTTCCCTCATCTCAGAGCCCTGGT 240
DB 299 GACAAAGGGGGCTTCTACATATCCCCCAAGACACTTTCAGCACTGTGAGAGCTGGT 358
QY 241 GACCAATTAC 249
DB 359 GACCACTAC 367

RESULT 12

BB619854 627 bp mRNA linear EST 31-AUG-2001
LOCUS BB619854
DEFINITION BB619854 RIKEN full-length enriched, adult male thymus Mus musculus
cDNA clone 5830437K10 5', mRNA sequence.
ACCESSION BB619854
VERSION BB619854.1 GI:15396929
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus

REFERENCE

1 (bases 1 to 627)
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
Arkawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A.
Hirose,K., Horii,F., Ishii,Y., Ito,M., Kawai,J., Kono,H., Kouda
M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K., Ono,M., Sakai
Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki
D., Shibata,K., Shingawa,A., Shiraki,F., Sogabe,Y., Suzuki,H.,
Tagami,M., Tagawa,A., Takahashi,F., Takeda,Y., Tanaka,T., Toya,T.,
Muramatsu,M. and Hayashizaki,Y.
RIKEN Mouse ESTs (Arkawa,T., et al. 2001)
Unpublished (2001)

TITLE

CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.lnl.gov
Plate: L1CM2052 row: 0 column: 13

Qy 241 GACATTAC 249
Db 255 CGCCATTAC 263

RESULT 14
LOCUS B1345674 591 bp mRNA linear EST 30-JUL-2001
DEFINITION B1345674 MARC 2P1G Sus scrofa cDNA 5', mRNA sequence.
ACCESSION B1345674
VERSION B1345674.1 GI:15038963
KEYWORDS EST.
SOURCE
ORGANISM Sus scrofa
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.
1 (bases 1 to 591)
Fahnenkrug, S.C., Freking, B.A., Rohrer, G.A., Smith, T.P.L., Casas, E.,
Stone, R.T., Heaton, M.P., Grosse, W.M., Bennett, G.A., Laegreid, W.M.
and Keefe, J.W.
Design and use of two pooled tissue normalized cDNA libraries for
EST discovery in swine
Unpublished (2000)
Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@email.marc.usda.gov
Single pass sequencing. Bases called and alt trimmed with phred
v0.980904.e. Vector identified by cross_match with the -minscore 18
and -minmatch 12 options.

TITLE
JOURNAL
COMMENT

FEATURES
source
1..591
/organism="Sus scrofa"
/db_xref="taxon:9823"
/clone_id="MARC 2P1G"
/tissue_type="pooled"
/lab_host="DH10B"
/note="Vector: PCMV SPORT6, Site 1: NotI; Site 2: SalI;
Library made from pooled tissue from testis, ovary,
endometrium, hypothalamus, pituitary, and placenta."
Location/Qualifiers

BASE COUNT 135 a 169 c 101 t
ORIGIN

Query Match 39.6%; Score 98.6; DB 13; Length 591;
Best Local Similarity 62.2%; Pred. No. 8.2e-17;
Matches 155; Conservative 0; Mismatches 94; Indels 0; Gaps 0;

Qy 1 TGCGTGTATGAGGCGCTGAGCAGGAGAAAGAGAGAACTGCTGTTTACCTGGGAAC 60
Db 178 TGCTTCTCAAGGCGATCAGCCGAGAGATGACAGCGCCAACTCTGCGCCCGGGAAC 237
Qy 61 CTTGAGAGGCGCTTCTCTCATCCGAGAGCCAGACAGAGAGAGCTTTACTCTGTCA 120
Db 238 GTGCTGGGCTCTTCATGATCCGGGACAGTGAGACACAAAGGAGACTTCTTTCG 297
Qy 121 GTCCGCTCAGCGCGCTGATCCTGGAGCGGATCGAGACATCAAGATCCACATGCTT 180
Db 298 GTGCGAGACTTACGAGCCCGAGACAGGAGACAGTGAGATTTACAAAGTCCGAGACCT 357
Qy 181 GACATGCTGCTGTACATCTACCGGCTTACCTTCCCTTCACTCCAGGCCCTGTG 240
Db 358 GACAGCGGGGCTTCTACATCTCCCAAGAGACTTCAACACCTGAGAGAGCTGTG 417
Qy 241 GACATTAC 249

Db 418 GCCCACTAC 426

RESULT 15
LOCUS BE268290 477 bp mRNA linear EST 13-JUL-2000
DEFINITION 601124314P1 NIH_MGC_8 Homo sapiens cDNA clone IMAGE:2989178 5',
mRNA sequence.
ACCESSION BE268290
VERSION BE268290.1 GI:9141893
KEYWORDS EST.
SOURCE
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Cetartiodactyla; Homnidae; Homo.
1 (bases 1 to 477)
NIH-MGC <http://mgi.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabs-r@mail.nih.gov
Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
cDNA Library Preparation: Ling Hong/Rubin Laboratory
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/BLNM, at: image.llnl.gov
Plate: L10M79 row: k column: 03
High quality sequence start: 35
High quality sequence stop: 474.
Location/Qualifiers

FEATURES
source
1..477
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_id="IMAGE:2989178"
/clone_lib="NIH MGC 8"
/tissue_type="Burkitt lymphoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: lymph. Vector: pOT8; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adapter: GGACGAG(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."

BASE COUNT 101 a 142 c 143 g 91 t
ORIGIN

Query Match 39.0%; Score 97; DB 10; Length 477;
Best Local Similarity 61.8%; Pred. No. 2.1e-16;
Matches 154; Conservative 0; Mismatches 95; Indels 0; Gaps 0;

Qy 1 TGCTGTATGAGGCGCTGAGCAGGAGAAAGAGAGAACTGCTGTTTACCTGGGAAC 60
Db 178 TGCTTCTCAAGGCGATCAGCCGAGAGATGACAGCGCCAACTCTGCGCCCGGGAAC 237
Qy 61 CTTGAGAGGCGCTTCTCTCATCCGAGAGCCAGACAGAGAGAGCTTTACTCTGTCA 120
Db 238 ACTCAGCGGCTCTTCTCATCCGAGAGAGAGAGACGCGGGATCGTTTCACTGTG 297
Qy 121 GTCCGCTCAGCGCGCTGATCCTGGAGCGGATCGAGACATCAAGATCCACATGCTT 180
Db 298 GTCCGAGACTTACGACGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 357
Qy 181 GACATGCTGCTGTACATCTTACCGGCTTACCTTCCCTTCACTCCAGGCCCTGTG 240
Db 358 GACAGCGGGGCTTCTACATCTCCCAAGAGACTTCAACACCTGAGAGAGCTGTG 417
Qy 241 GACATTAC 249
Db 418 CGCCATTAC 426

Tue Apr 1 06:01:48 2003

us-09-988-971-1_copy_694_942.rst

Page 10

Search completed: March 30, 2003, 02:49:25
Job time : 494.002 secs
